A TRIDENT SCHOLAR
PROJECT REPORT

NO. 266

NON-INVASIVE DETECTION OF CH-46 AFT GEARBOX FAULTS USING
DIGITAL PATTERN RECOGNITION AND CLASSIFICATION TECHNIQUES

UNITED STATES NAVAL ACADEMY
ANNAPOLIS, MARYLAND
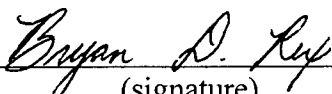
20000424 157

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including g the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 5 May 1999	3. REPORT TYPE AND DATE COVERED

4. TITLE AND SUBTITLE
Non-invasive detection of CH-46 AFT gearbox faults using digital pattern recognition and classification techniques

6. AUTHOR(S)
Rex, Bryan D.

5. FUNDING NUMBERS

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Naval Academy
Annapolis, MD

8. PERFORMING ORGANIZATION REPORT NUMBER
USNA Trident Scholar project report no. 266 (1999)

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Accepted by the U.S. Trident Scholar Committee

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document has been approved for public release; its distribution is UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT: Currently, the United States Navy performs routine intrusive maintenance on CH-46 helicopter gearboxes in order to diagnose and correct possible fault conditions (incipient faults) which could eventually lead to gearbox failure. This type of preventative maintenance is costly and it decreases mission readiness by temporarily grounding usable helicopters. Non-invasive detection of these fault conditions would save time and prove cost-effective in both manpower and materials. This research deals with the development of a non-invasive fault detector through a combination of digital signal processing and artificial neural network (ANN) technology. The detector will classify incipient faults based on real-time vibration data taken from the gearbox itself. Neural networks are systems of interconnected units that are trained to compute a specific output as a non-linear function of their inputs. For some time the United States Navy has been interested in the use of artificial neural networks in monitoring the health of helicopter gearboxes. In order to determine the detection sensitivity of this method in comparison with traditional invasive methods, the USN funded Westland Helicopters Ltd to conduct a series of CH-46 gearbox rig tests. In these tests, the gearbox was seeded with nine different fault conditions. This seeded fault testing provided the vibration data necessary to develop and test the feasibility of an artificial neural network for fault classification. This research deals with the formation of the pattern vectors to be used in the neural network classifier, the construction of the classification network, and an analysis of results.

14. SUBJECT TERMS
Artificial neural networks; condition based maintenance, digital signal processing; fault diagnostics; health monitoring; incipient faults; pattern recognition

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

NON-INVASIVE DETECTION OF CH-46 AFT GEARBOX FAULTS USING DIGITAL PATTERN RECOGNITION AND CLASSIFICATION TECHNIQUES
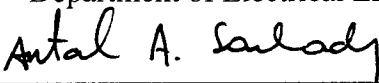
by

Midshipman Bryan D. Rex, Class of 1999
United States Naval Academy
Annapolis, Maryland
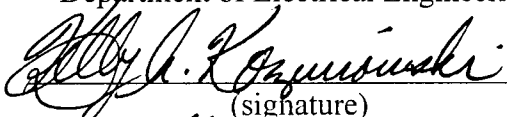
_____
(signature)

Certification of Adviser(s) Approval

Professor Antal A. Sarkady
Department of Electrical Engineering

_____
(signature)

May 4, 1999
(date)

Associate Professor Kelly A. Korzeniowski
Department of Electrical Engineering

_____
(signature)

May 5, 1999
(date)

Acceptance for the Trident Scholar Committee

Professor Joyce E. Shade
Chair, Trident Scholar Committee

_____
(signature)

5 May 1999
(date)

USNA-1531-2

Abstract

Currently, the United States Navy performs routine intrusive maintenance on CH-46 helicopter gearboxes in order to diagnose and correct possible fault conditions (incipient faults) which could eventually lead to gearbox failure. This type of preventative maintenance is costly and it decreases mission readiness by temporarily grounding usable helicopters. Non-invasive detection of these fault conditions would save time and prove cost-effective in both manpower and materials. This research deals with the development of a non-invasive fault detector through a combination of digital signal processing and artificial neural network (ANN) technology. The detector will classify incipient faults based on real-time vibration data taken from the gearbox itself.

Neural networks are systems of interconnected units that are trained to compute a specific output as a non-linear function of their inputs. For some time the United States Navy has been interested in the use of artificial neural networks in monitoring the health of helicopter gearboxes. In order to determine the detection sensitivity of this method in comparison with traditional invasive methods, the USN funded Westland Helicopters Ltd to conduct a series of CH-46 gearbox rig tests. In these tests, the gearbox was seeded with nine different fault conditions. This seeded fault testing provided the vibration data necessary to develop and test the feasibility of an artificial neural network for fault classification. This research deals with the formation of the pattern vectors to be used in the neural network classifier, the construction of the classification network, and an analysis of results.

Key Words: Artificial neural networks; condition based maintenance, digital signal processing; fault diagnostics; health monitoring; incipient faults; pattern recognition

Acknowledgments

I would like to express my heartfelt gratitude to everyone whose encouragement and support have guided me through this project. First, I would like to thank God for providing me with the perseverance and reasoning to accomplish what I have. Mom and Dad, thanks for always being there and pushing me when it was needed. Daphi, your computer expertise and numerous hours spent fixing the lab computers and printers did not go unnoticed. Lastly, I would like to give my sincere thanks to my advisors, Professor Antal A. Sarkady and Associate Professor Kelly A. Korzeniowski, who have supported me throughout the entire process.

Contents

Introduction

Traditionally, the Navy has used invasive methods for preventative maintenance of helicopter gearboxes. These methods have proven costly in both manpower and resources. A requirement to improve several aspects of fault detection has existed for several years, especially within the rotary wing community. These requirements have been set forth to improve mission readiness through more effective maintenance, elimination of losses of aircraft and personnel, and reduction of maintenance related costs [4]. In addition, the need to extend operational service lifetimes of aircraft as well as a reduction of manpower have made these improvements more urgent. The use of non-invasive diagnostic procedures allows aircraft faults to be diagnosed at the organizational level (during normal service), as opposed to discovery during tear-down at the intermediate or depot level. Depot level includes rework facilities such as the Naval Air Rework Facility at Cherry Point.

This research involves the timely detection of CH-46 helicopter gearbox faults through non-invasive vibration monitoring. An example of a typical CH-46 Helicopter mission is illustrated in Figure a. Digital signal processing coupled with a pattern recognition algorithm, such as an artificial neural network or a Bayesian network, provides a promising means of classifying real-time vibration data for fault detection. Several methods of fault detection for rotary winged aircraft are currently used by the United States Navy, but none have proven 100 percent effective at preventing catastrophic failure, and most cannot specifically identify drive-train faults

[4]. These methods include the use of chip detectors, the Navy Oil Analysis Program (NOAP) [15], component cards, and vibration analysis. More recently, the development of the Helicopter Integrated Diagnostic System (HIDS) [4], and the use of commercial-off-the-shelf (COTS) components have

Fig. a: CH-46 Performing Vertical Replenishment

vastly improved the ability to perform condition based maintenance.

Recently, Westland Helicopter Ltd. collected the vibration data necessary to investigate the possibility of applying new methods for determining incipient fault conditions. The data, collected by Westland Helicopters Ltd. and digitized by NRAD (Naval Research and Development Center) in San Diego, California, include representative vibration characteristics of a CH-46 gearbox under several different conditions (both defective and non-defective). The conditions tested include eight specific fault areas which are listed below:

- no defect
- input pinion bearing corrosion (first and second defect level)
- spiral bevel input pinion spalling (first and second defect level) (Fig. b)
- helical input pinion chipping (second defect level)
- collector gear cracking
- quill shaft cracking

- planetary bearing corrosion
- helical idler gear cracking.
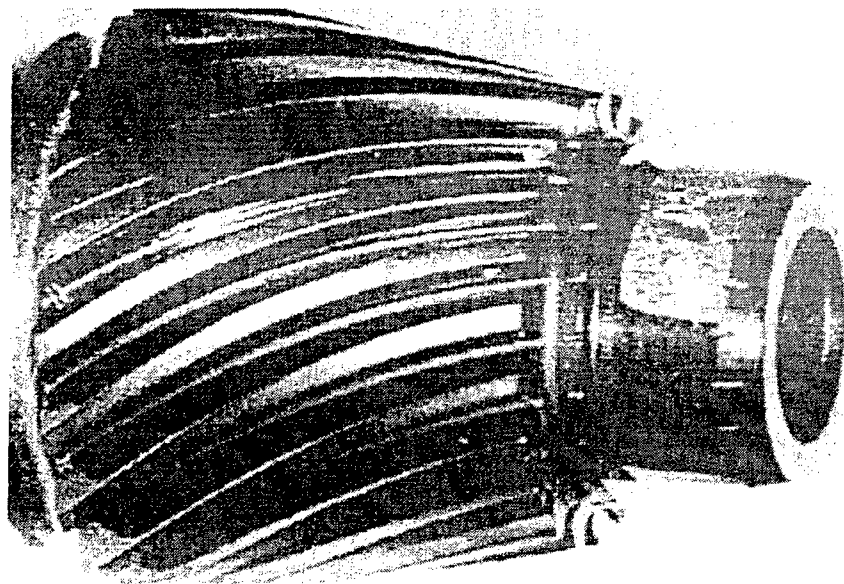
Fig. b: Input Pinion Spalling (NAWC) |4|

A single mixbox and one aft main transmission were installed on a test rig (Fig. c) and run at nine different torque conditions. Vibration data were recorded using eight different accelerometers and an optical tachometer with an analog tape recorder. Only one faulty component at a time was introduced into the gearbox during each of the test runs. Each of the test runs was conducted over a sufficient period of time to provide reproducible and representative gearbox vibration information. The test rig used to monitor these conditions provided a safe means of collecting data that - if encountered during normal operation - could lead to tragedy.
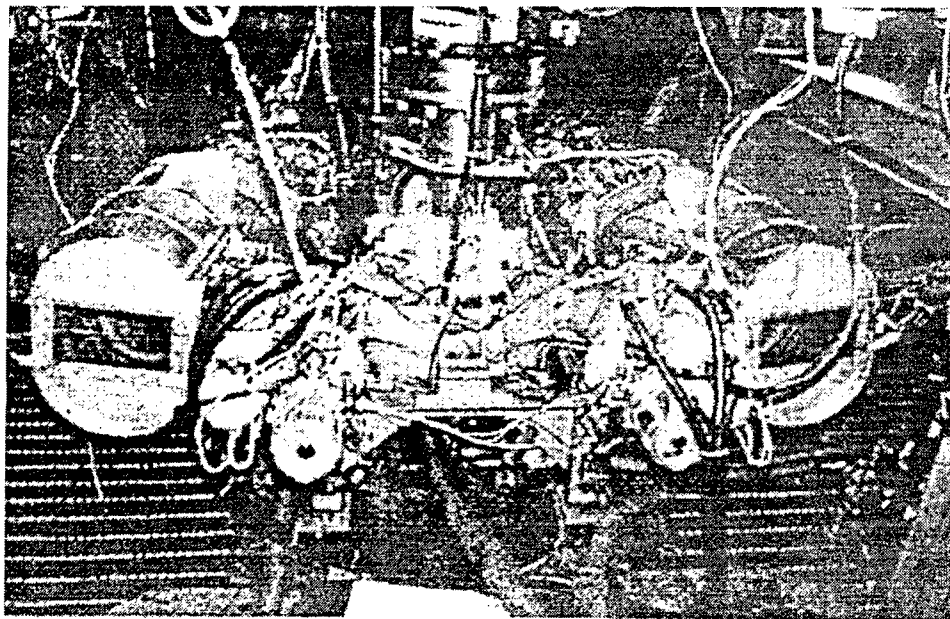
The **Fig. c: General Test Rig Assembly (NAWC) [4]** obje

ctive of this research is to develop the digital signal processing and classification

techniques necessary to implement non-invasive fault testing on an aft CH-46

gearbox. One of the primary aims of this project is reduction of the data set by

determining the important signal characteristics and filtering out the unnecessary data.

By determining what characterizes each individual flaw in the Westland data set, a

more general "fingerprint" can be established so that similar flaws in other rotating

machines can also be detected. This project not only provides a method for detecting

these specific fault conditions in a CH-46 gearbox, but furnishes the groundwork for

applying this method of fault detection to other rotating devices sharing similar

components.

Chapter 1:
Data Collection and the CH-46 Aft Gearbox

The main problem with creating a reliable, on-board diagnostics system has been the lack of raw data needed to characterize fault conditions. Since most Class A mishaps (loss of aircraft and/or personnel) are due to engine/drive-train failures, the need to collect information about engine/drive-train faults is crucial.

1.1 The Westland Universal Test Rig

Westland's universal transmission test rig was intended for fatigue testing of helicopter gearboxes with up to three driving inputs and a single output, and is composed of 3 x 3500 shaft horsepower electric drives (capable of 25000 rpm), and two water brake dynamometers capable of absorbing up to 6000 shaft horsepower. [2] The 'Magna Power' electronic drives were coupled to the gearbox through an overdrive gear system coupled to a high speed reversing gearbox. Since the helical input pinion on a CH-46 turns at 324.60 Hz during normal operation, the electronic drive's shaft frequency of 49.95 Hz was stepped up to 324.60 Hz by an overdrive stage. The schematic for the test rig is illustrated in Figure 1.1.

1.2 Instrumentation

The instrument package used to monitor gearbox vibration was supplied by the Naval Air Warfare Center (NAWC), Aircraft Division (Patuxent River). The
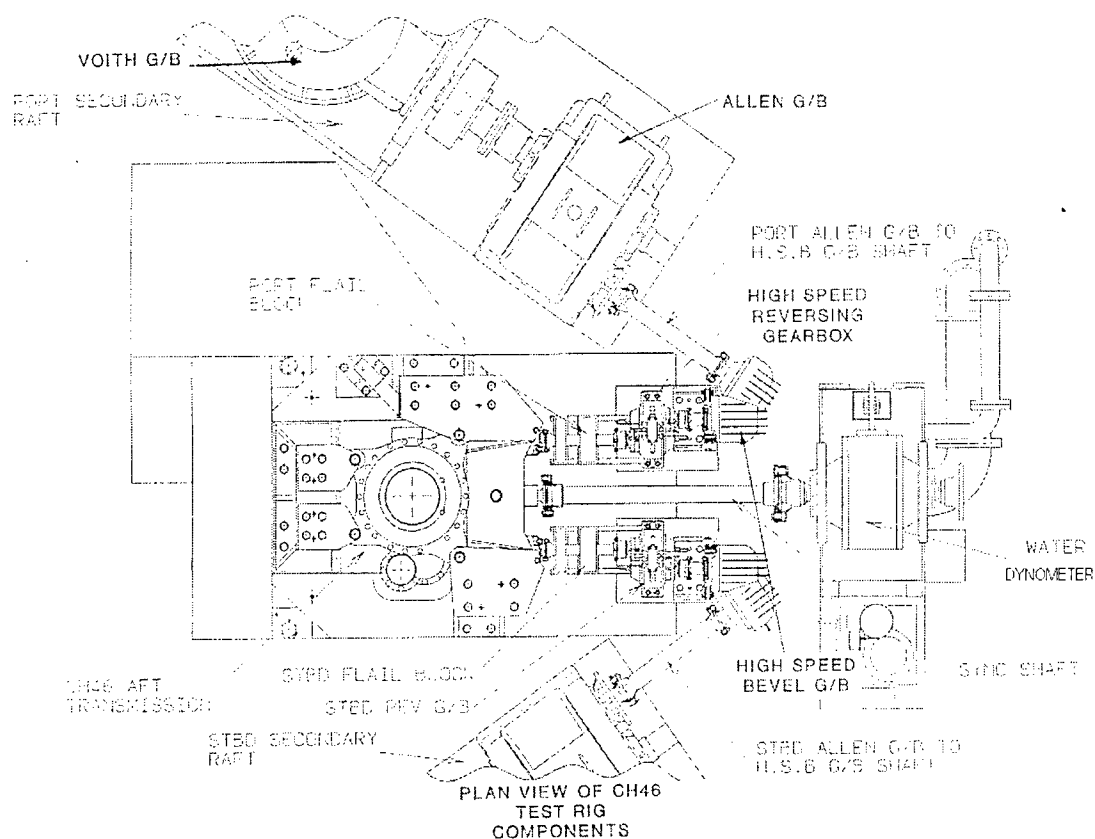
Fig. 1.1: Test Rig and Components (Westland Helicopters) [2]

package included eight 'Endevco 7259A' accelerometers. which were mounted on

special brackets also supplied by the NAWC. Also placed on the gearbox was an

optical tachometer that fitted in place of the blade fold drive motor. The inputs from

each of the eight accelerometers, the tachometer signal, and a tape servo reference

tone were all recorded on individual channels of a 28 channel 'Racal Storehouse'

analog tape recorder at a rate of 15 inches per second. The analog information was

later filtered with a non-aliasing filter and digitized at a sampling rate of 103,116.08

Hz. Figure 1.2 shows the gearbox and sensor placement (sensors 4,5,6, and 8).
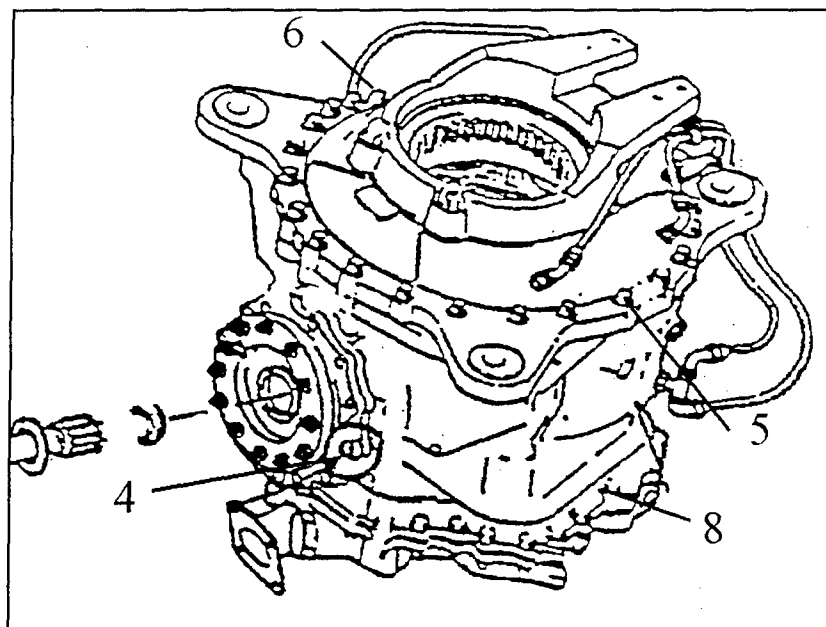
Fig. 1.2: Sensor Placement (Westland Helicopters) [2]

1.3 The CH-46 Aft Gearbox

An important part of analyzing the data collected by Westland Helicopters involves understanding the basic operation of the gearbox itself. Primarily, the gear mesh frequencies (the product of shaft frequency and number of gear teeth), shaft frequencies, and resonant frequencies of internal parts can correlate with the vibration characteristics associated with specific fault conditions. Figure 1.3 illustrates the basic schematic for a CH-46 aft gearbox. The input shaft frequency (324.60 Hz) is reduced by a helical idler gear to 126.23 Hz. The shaft speed is then further reduced by a spur pinion/collector gear to 42.65 Hz (the collector gear combines the port and starboard inputs). The quill shaft is driven by the collector gear at 42.65 Hz, and its speed is again reduced by a spiral bevel pinion/gear combination. The spiral bevel gear turns at 17.60 Hz, and its shaft frequency is further reduced to 4.40 Hz through

yet another reduction stage involving a sun, planetary, and ring gear combination [2].

Table 1.1 shows shaft and gear mesh frequencies for the main gearbox parts. Figure

1.3 illustrates the gearbox schematic.
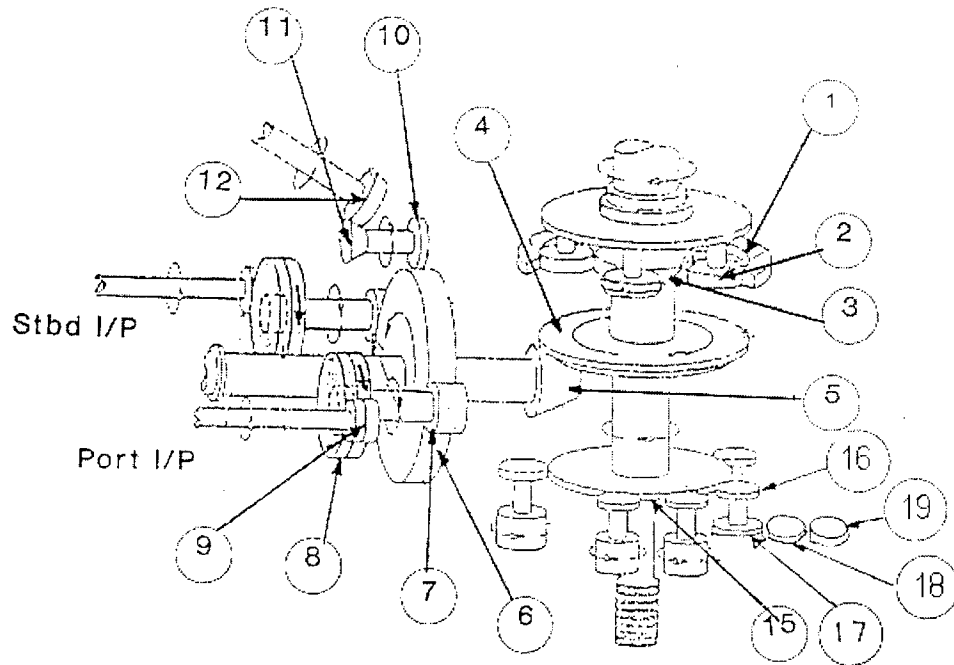
Figure 1.3: CH-46 Gearbox Schematic (Westland Helicopters) [2]

Table 1.1: Shaft and Gear Mesh Frequencies

Part	Shaft Frequency	No. of Teeth	Gear Mesh Freq.
Helical Input Pinion (9)	324.60 Hz	28	9088.8 Hz
Helical Idler Gear (8)	126.23 Hz	72	9088.8 Hz
Spur Pinion (7)	126.23 Hz	25	3155.75 Hz
Collector Gear (6)	42.65 Hz	74	3155.75 Hz
Blower Spur Pinion (10)	126.23 Hz	25	3155.75 Hz
Blower Bevel Gear (11)	126.23 Hz	25	3155.75 Hz

Part	Shaft Frequency	No. of Teeth	Gear Mesh Freq.
Blower Pinion (12)	101.80 Hz	31	3155.75 Hz
Quill Shaft	42.65 Hz	-	-
Spiral Bevel Pinion (5)	42.65 Hz	26	1108.90 Hz
Spiral Bevel Gear (4)	17.60 Hz	63	1108.90 Hz
Sun Gear (3)	-	39	514.80 Hz
Planet Gear (2)	-	39	514.80 Hz
Ring Gear (1)	-	117	514.80 Hz
Aux Drive Gear (15)		130	2288 Hz
Accessory Drive Aft (16)		20	2288 Hz
Rotor Position Drive (17)		42	4804.8 Hz
Scavenge Pump Drive (18)		64	4804.8 Hz
Optical Tach Drive (19)		42	4804.8 Hz
Output Shaft	4.40 Hz	-	-

Chapter 2:

Digital Signal Processing

2.1 Initial Processing of Raw Data

The initial step in processing the raw data involved reading the digitized

information into the computer so that it could be processed and manipulated. The

data were digitized by NRAD at a sample rate of 103,116.08 Hz with 16 bit

quantization using a ten channel data acquisition system. The data format was 16-bit

two's complement (short integer, big-endian). It was sample multiplexed into 20-byte

frames on the CDs. The multiplex scheme is shown below in Table 2.1:

Table 2.1: Data storage scheme for digitized vibration data

Bytes 1-2	Channel 1	800Hz Reference Tone
Bytes 3-4	Channel 2	Accelerometer #1
Bytes 5-6	Channel 3	Accelerometer #2
Bytes 7-8	Channel 4	Accelerometer #3
Bytes 9-10	Channel 5	Accelerometer #4
Bytes 11-12	Channel 6	Accelerometer #5
Bytes 13-14	Channel 7	Accelerometer #6
Bytes 15-16	Channel 8	Accelerometer #7
Bytes 17-18	Channel 9	Accelerometer #8
Bytes 19-20	Channel 10	Tachometer

One file of this format contained approximately 21 seconds of data taken at

normal operating speed and temperature. Each of these files contained data

corresponding to each of the eight sensors for a specific fault at set torque level. Before any type of signal processing could be performed on the data, it was necessary to de-multiplex the data into files that contained data for each individual sensor. This was accomplished by writing a MATLAB "m" file that parsed the data for the accelerometers, the reference tone, and the tachometer into ten individual files. This program was applied to each of the 68 original files, creating 680 output files of 3.9 Mb in size. These smaller files were then saved to the computer's hard drive.

2.2 Digital Demodulation

The operation of any gearbox centers around the rotation of the shafts and gears that compose the machine. A non-faulted gearbox would tend to be balanced and function more smoothly than one with a fault condition present. A cracked shaft or gear would cause vibrations that are superimposed on the normal rotational vibrations. Intuitively, this can be viewed as a modulation process [12]. It was hypothesized that amplitude, phase, and frequency modulation (AM, PM, and FM) would be apparent in the accelerometer signals. In order to take advantage of this characteristic, the analytic signal (defined below) was formed and used to calculate the envelope and phase of the original signals. The digital demodulation process provided a means to reduce the original data set greatly.

The Hilbert transform was the first step in forming the analytic signal. For a real signal $f(t)$, the Hilbert transform [14] is defined in the time domain (denote by $\otimes$)

as the convolution of $f(t)$ by $1/(\pi \cdot t)$ as defined by Equation (2.1).

$$f_h(t) = f(t) \otimes \frac{1}{\pi \cdot t} = \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} \frac{f(\tau)}{t - \tau} d\tau \qquad (2.1)$$

To express the Hilbert transform function. $f_h(t)$, in the frequency domain, we apply the Fourier transform to Eq. 2.1, where $\Im\{\}$ denotes the continuous Fourier transform operator [14] and $F(\omega)$ is the Fourier spectrum of the original signal.

$$\Im\{f_h(t)\} = F_h(\omega) = F(\omega) \cdot \Im\{\frac{1}{\pi \cdot t}\} \qquad (2.2)$$

where $\Im\{\frac{1}{\pi \cdot t}\} = \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} \frac{1}{t} e^{-j \cdot \omega \cdot t} dt = -j \cdot \text{sgn}(\omega)$ \qquad (2.3)

$sgn(\omega)=1$ for $\omega>0$, 0 for $\omega=0$, and -1 for $\omega<0$.

Therefore, $\qquad F_h(\omega) = (-j \cdot sgn(\omega)) \cdot F(\omega)$. \qquad (2.4)

$\qquad\qquad F_h(\omega) = -j \cdot F(\omega)$ for $\omega<0$ \qquad (2.5)

$\qquad\qquad\qquad = j \cdot F(\omega)$ for $\omega>0$

$\qquad\qquad\qquad = 0 \qquad$ for $\omega=0$

Now that the Hilbert Transform is defined, the analytic signal representation of $f(t)$ can be easily determined. The analytic signal representation of $f(t)$ is a complex valued signal in the time domain with a one sided spectral density in the frequency

domain [14]. The real part of the analytic signal, $z(t)$, is equal to the original signal. The imaginary part of $z(t)$ is equal to the Hilbert Transform of $f(t)$. This relationship can be written mathematically as:

$$z(t) = f(t) + jf_h(t) \tag{2.6}$$

By taking the Fourier transform of $z(t)$, we find:

$$Z(\omega) = F(\omega) + jF_h(\omega) \quad = \quad F(\omega) + j(-j\,sgn(\omega) \cdot F(\omega)) \tag{2.7}$$

From the definition of $F_h(\omega)$ above (eqn 2.5), it can be verified that the spectral density of $z(t)$ is a one-sided function in the frequency domain:

$$
\begin{aligned}
Z(\omega) \quad &= \quad 2F(\omega) \quad \text{for } w > 0 \\
&= \quad F(\omega) \quad \text{for } w = 0 \\
&= \quad 0 \quad\quad \text{for } w < 0
\end{aligned}
\tag{2.8}
$$

In other words, $Z(\omega)$ is an upper single-sideband signal in the baseband which can be found by doubling the positive side of the original frequency spectrum, and zeroing its negative components. In Discrete Fourier Transform (DFT) sense, the negative side of the frequency spectrum lies between $N/2+1$ and $N-1$, where N is the number of points used with the DFT. A discrete version of the analytic signal $z[n]$ can be determined in the time domain by taking its DFT, doubling the positive spectrum, and then taking its Inverse DFT (IDFT).

An important application of the analytic signal, $z(t)$, is that it can be used for demodulation when $f(t)$ can be modeled as an amplitude, frequency, or phase modulated function (AM, FM, or PM) [14]. If $f(t)$ is a double side-band with large-

carrier AM signal (DSB-LC), it can be shown that the absolute value of $z(t)$, or the

envelope, recovers the modulating component of the AM signal. In addition, the

phase of $z(t)$ can be used to recover the modulating component when $f(t)$ is modeled

as an FM or a PM signal. It is necessary to remove the discontinuities in the

computational process of the phase function by using the "unwrap" command in

MATLAB. This command removes the computational discontinuities in the radian

phase by changing absolute phase jumps greater than pi to their 2*pi complement. A

linear regression algorithm is applied to the unwrapped signal, and the straight line

carrier trend is computed and subtracted. The remaining difference phase signal is

defined as the demodulated phase function which was accompanying the carrier trend.

This demodulated signal is in the base-band and is referred to in communication

theory as the angle modulation on the carrier. This modulation can be attributed to

either phase or frequency change. If phase modulation is assumed, then the signal is

used directly. If frequency modulation is assumed, then the derivative of the angle

modulation provides the frequency modulation (FM). Therefore the formation of the

analytic signal provides a means to AM, PM, and FM demodulate the original signal.

In order to apply this demodulation technique on a finite length signal, the following

algorithm steps must be employed:

- Take the FFT of the signal

- Apply the analytic signal filter in the frequency domain (as defined above)

- Compute the phase of the analytic signal

- Unwrap the phase function

- Compute and subtract the linear carrier trend (producing the demodulated PM signal)

- Differentiate the PM signal in order to produce the demodulated FM signal.

2.3 Ensemble Averaging

It is not uncommon for a real information carrying signal to be masked in additive noise, such as random Gaussian noise. Depending on the signal-to-noise voltage ratio (which is defined as the ratio of the root mean square of the signal to the root mean square of the noise), a single Fourier spectral estimate may be sufficient to identify and quantify the spectral lines in the computed spectrum. If the signal to noise ratio is poor, then the process of ensemble averaging can help in the identification of spectral lines. In the ensemble averaging process (or Bartlett smoothing procedure [5]), the original signal is windowed in the time domain as described below. The Fourier magnitude spectrum is then calculated for each of the independent records. Since the phase information is lost in this transformation, the spectral estimates can be averaged, providing a statistically reliable frequency spectrum whose signal to noise ratio is improved approximately by the square root of the number of records averaged when the number of records is large (greater than 100). This process is very important in reducing the number of points in the data set, while preserving useful information. A long record may consist of several million points, while the resultant ensemble frequency spectrum may consist of only a few hundred points.

Each of the independent time records was windowed in the time domain by taking the product of a Hanning window function with the segmented record. The windowing reduces spectral leakage at the expense of frequency resolution. In general, a Hanning (or cosine-squared) window function is defined as [14]:

$$w(t) = \cos^2 \frac{\pi \cdot t}{\tau} = \frac{1}{2}(1 + \cos\frac{2\pi t}{\tau}) \text{ for } |t| \leq \frac{\tau}{2} \text{ and } w(t) = 0 \text{ elsewhere} \quad (2.10)$$

The coefficients of the Hanning window are determined by:

$$w(m) = \frac{1}{2}(1 + \cos\frac{m\pi t}{10}) \text{ for } |m| \leq 10 \quad (2.11)$$

The use of a Hanning window results in a frequency spectrum whose frequency resolution is decreased by a factor of two over that of a standard rectangular window, yet reduces the nearest leakage lobe by approximately 16 dB. For example, an N-point FFT of rectangular windowed data will have twice the frequency resolution of an N-point FFT of data windowed using a Hanning algorithm. The frequency spectrum of each of the windowed time records was then calculated, and the resultant frequency spectra were then averaged in order to produce one spectrum whose signal-to-noise ratio was improved by a factor of the square root of the number of averaged records. The ensemble average frequency (magnitude) spectrum was calculated for the signals from each sensor for every fault condition. Sampled amplitudes of frequency components will serve as inputs to the neural network to be

used for classification later. During the ensemble process. the standard deviation (root mean square value) of the individual time records was also determined. This will also be used as an input to the neural network.

The ensemble averaging technique was also used to determine the frequency spectrum of the envelope, PM, and FM signals. Again, several independent time records were demodulated using the analytic signal. The frequency spectra of the AM, PM, and FM signals were then found and ensemble averaged as explained above. The root mean square, RMS, value of the demodulated signals was also determined by averaging the RMS value among each of the separate time records for each envelope, PM, and FM signal. This process was carried out for the signals from each sensor for every fault condition.

2.4 Peak Detection, Moving Average Filter and Signal-to-Noise Ratio

Although it was easy to distinguish rugged signal characteristics visually from the frequency spectra, the sheer amount of data present made it necessary to automate the process. In order to determine rugged features, a peak detection filter was developed that compared the area beneath a signal peak to the area of the surrounding noise. If this ratio was above a specified threshold, then the point was deemed rugged and kept for formation of the pattern vector. This filter also provided a means of determining the signal-to-noise ratio, which can also be used to help identify gearbox condition. Since it is expected that the background noise will change during gearbox operation, determination of the signal-to-noise ratio is a suitable method to

Figure 2.1 Continuum in Frequency Spectrum
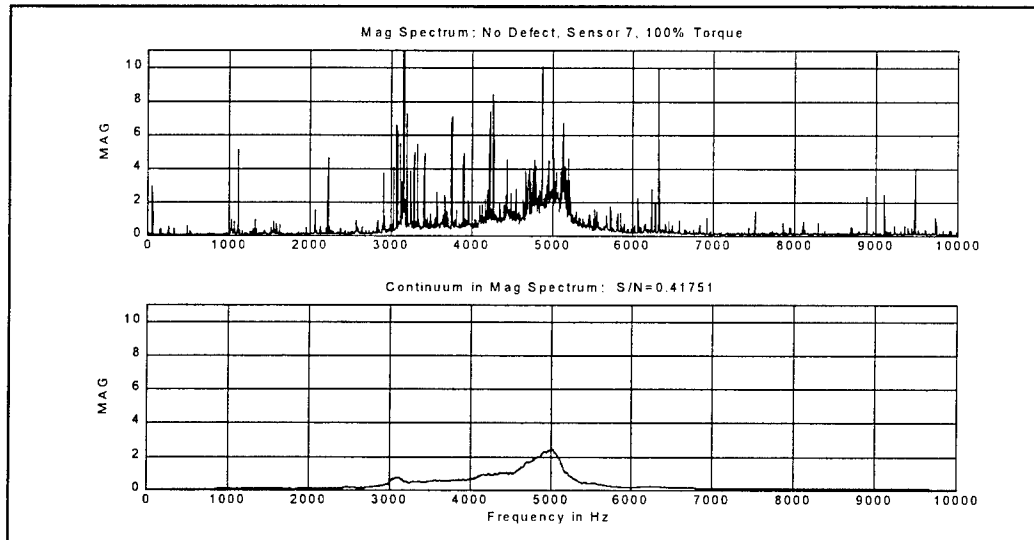
describe important features quantitatively. By subtracting the signal peaks from the original signal and applying a fiftieth order moving average filter, a broad-band continuum was determined that provided yet more insight into gearbox health. This continuum is a result of either noise, short-duration impulses in the time-domain, or a combination of both. Regardless, each flaw had a unique continuum associated with it. An example of such a continuum is illustrated in Figure 2.1.

2.5 Processing the Data

The primary goal involved in this step was to determine important, or "rugged", signal characteristics that describe the individual fault conditions. These features would provide a means of distinguishing a faulty gearbox from a good gearbox, and a means of pinpointing the actual flaw if one were to occur. The

techniques described earlier served as tools used to determine the important signal characteristics. The necessary steps during the processing were:

- to determine and examine the frequency spectra of the signals from the eight accelerometers corresponding to each of the fault conditions
- to find the RMS value of each of these signals
- to use the analytic signal representation of each of the signals in order to perform the AM, PM, and FM demodulation
- to determine and examine the frequency spectra of the envelope, PM, and FM signals derived from each of the original signals
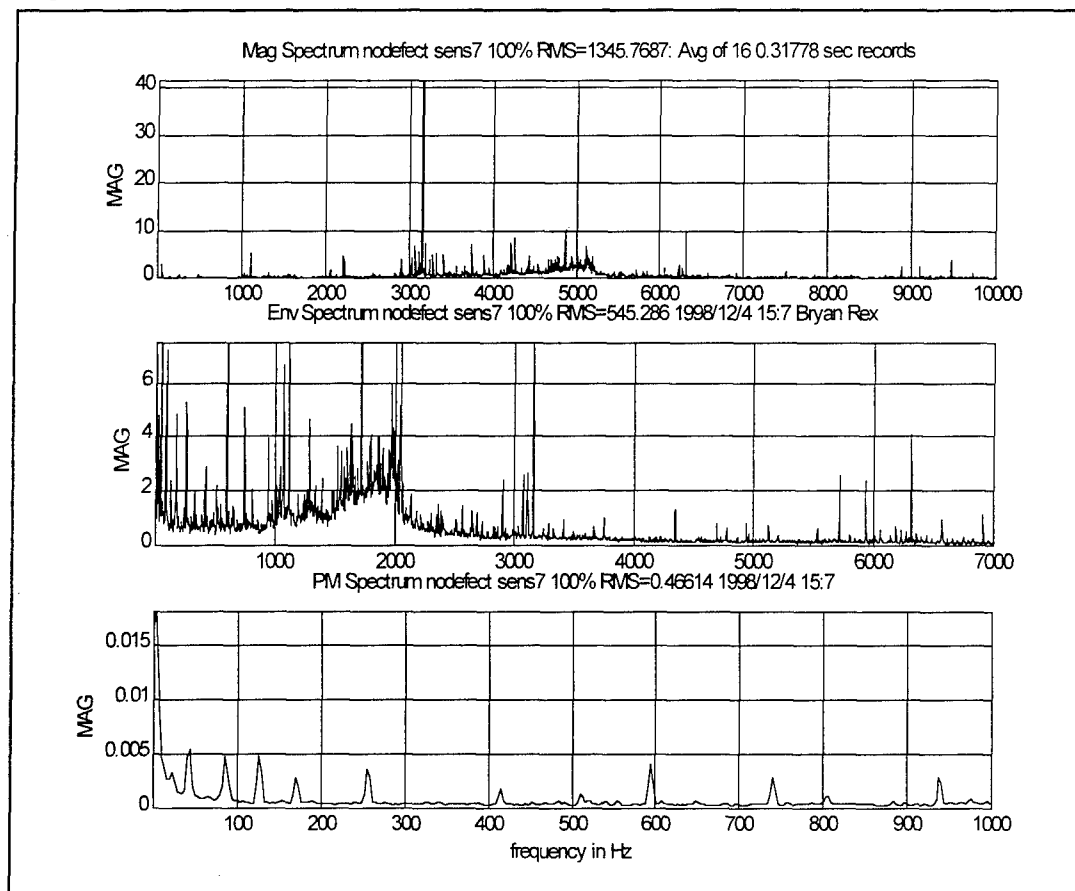- to find the RMS value of each envelope, PM, and FM signal

The trade-off between frequency resolution and signal-to-noise ratio was a major consideration when determining the number of points to use in the FFT. Since approximately one-third of the data will be used to train the neural network classifier, 6 seconds (618,696 points) of data were initially processed for each individual signal (each of the sampled signals comprise nearly 21 seconds of raw data). With a sampling rate of 103,116.08 Hz, a single 618,696 point FFT would have provided 1/3 Hz frequency resolution (using a Hanning window), yet the signal-to-noise ratio would remain unimproved since only one record would be used in the ensemble average. Since the aim was to increase the signal-to-noise ratio as much as possible while retaining frequency resolution, the number of points used in the FFT was gradually increased until all line-splitting ceased. It was determined that line splitting stopped when frequency resolution was roughly 7 Hz. With a Hanning window algorithm, 7 Hz resolution is achieved when the number of points (n) used in the FFT is $n>2*fs/7$. Since fs=103,116.08, the number of points needed in the FFT was at

least 29,461. By using a value of n that is a power of 2, the FFT can be employed as

opposed to the DFT. The number of complex operations involved in the FFT is equal

to $n\dfrac{\log(n)}{\log(2)}$ while the DFT involves n^2 complex operations [13]. Therefore, by

using n=2^{15}=32768 points as oppposed to 29461, the processing time is sped up by a

factor of approximately 1,766. Therefore, the original 618,696 point records were

divided into independent 32.768 point records and windowed using the Hanning

algorithm. Further inspection indicated that the signal-to-noise ratio was fairly high,

therefore many ensemble averages were not necessary. It was decided that improving

the signal-to-noise ratio by a factor of four was more than sufficient to identify

rugged features, therefore 16 independent records were ensemble averaged .

Once it was decided that the data needed to be divided into mutually exclusive

32,768 point (.318 second) records, a major part of this research involved writing a

program in MATLAB that performed the calculations necessary to determine and

save the following information into files: the ensemble frequency spectrum, the RMS

value, the ensemble envelope spectrum (AM spectrum), AM RMS value, the

ensemble PM spectrum, the PM RMS value, the ensemble FM spectrum, and the FM

RMS value for each of the 580 original signals. The program output this data

graphically, saved the vectors to files, and also print the graphical output which

consist of the frequency spectrum, the envelope spectrum, the PM spectrum, along

with the RMS value of each of these signals. This program reduced the number of

Fig. 2.2: Output for Non-Faulted Condition

Mag Spectrum nodefect sens7 100% RMS=1345.7687: Avg of 16 0.31778 sec records

Env Spectrum nodefect sens7 100% RMS=545.286 1998/12/4 15:7 Bryan Rex

PM Spectrum nodefect sens7 100% RMS=0.46614 1998/12/4 15:7

points in the data set by approximately 96%. Examples of the graphical output for

the no defect condition at sensor 7 (Fig 2.2) is compared to the output for the quill

shaft crack condition (Fig 2.3).

The helical input pinion shaft turns at a constant rate of 324.60 Hz during

gearbox operation. Therefore frequency normalization is unnecessary, yet can be

achieved by dividing the frequency index of the magnitude spectrum by the frequency

of the tachometer signal.

The first plot both Fig. 2.2 and Fig. 2.3 represents the ensemble averaged

magnitude spectrum for each condition. Sixteen individual .318 second time records

Figure 2.3 Output for Quill Shaft Crack Condition

Mag Spectrum qsc sens7 100% RMS=1107.5284: Avg of 16 0.31778 sec records

Env Spectrum qsc sens7 100% RMS=600.3525 1998/12/5 4:25 Bryan Rex

PM Spectrum qsc sens7 100% RMS=10.5199 1998/12/5 4:25
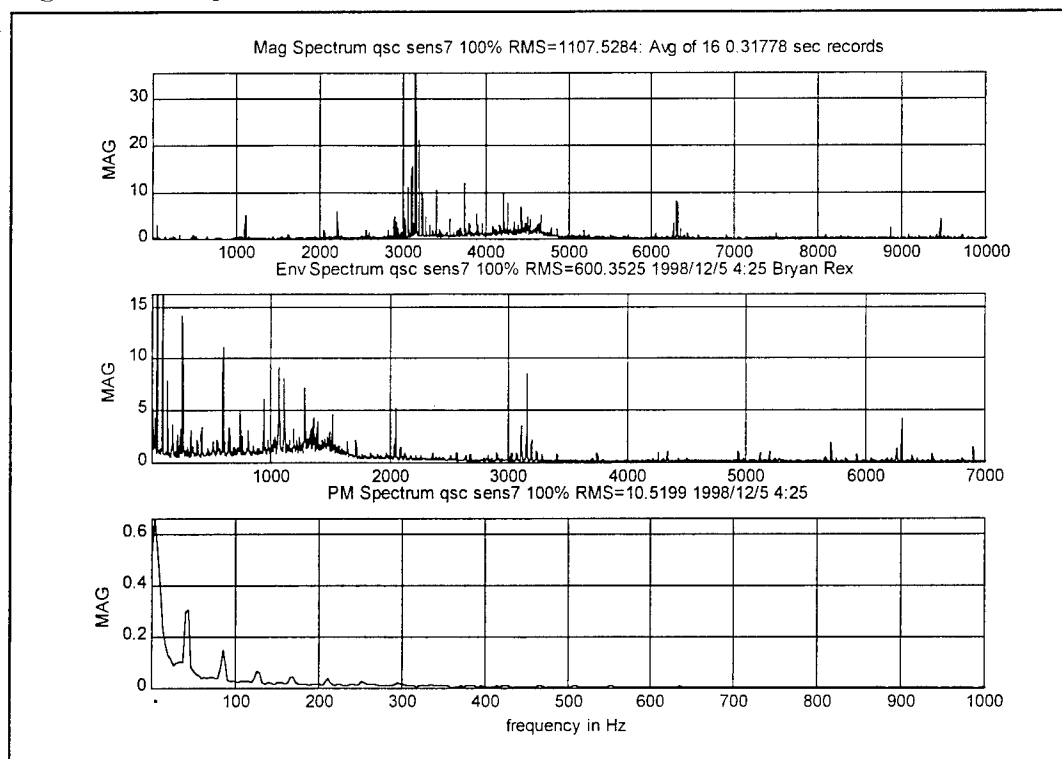
frequency in Hz

were averaged in order to achieve 6.3 Hz resolution and an improved signal-to-noise ratio. It becomes visually apparent that the quill shaft fault suffers frequency modulation, with carrier frequencies of 3156, 6312, and 9468 Hz. The line structure (FM modulation) that appears in this defect around the 3156 Hz and 6312 Hz peaks of the magnitude spectrum became apparent only by increasing the frequency resolution of the FFT to at least 12.6 Hz (16,384 points using a Hanning window). This information would be lost without using at least a 16,384 point FFT. The large amplitude peak at 3156 Hz is common to both conditions, as well as the harmonically related peaks at 6312 Hz, and 9468 Hz. These harmonics are hypothesized to be a result of vibration in the spur pinion, collector gear, and blower bevel pinion/gear, which all have mesh frequencies of 3155.75 Hz.

The second plot in each figure represents the spectrum of the envelope associated with each fault. This was found using the analytic signal and complex demodulation technique described earlier. As with the magnitude spectrum, sixteen individual .318 second records were averaged in order to improve signal-to-noise ratio. The quill shaft crack shows signs of frequency modulation, with a modulating frequency of about 43 Hz. It is also apparent that the continuum associated with the quill shaft crack occupies a narrower bandwidth than that of the non-faulted condition.

The third plot in each figure is the spectrum of the phase modulation. The phase modulation and its derivative, frequency modulation, were also found using the analytic signal. Again, the quill shaft fault shows a fairly large peak at 43 Hz (note that the y-axes differ between the plots). This is consistent with the fact that the quill shaft turns at 43 Hz within the gearbox. As hypothesized, the crack in the shaft seems to be causing a modulation at the shaft frequency.

Also apparent is a unique continuum in both frequency spectra (as illustrated in Fig 2.1). It is hypothesized that this continuum is due partly to noise, and partly to short impulse-like events in the time domain. The duration of such an event can be determined by $T_{event}=1/$Bandwidth of the continuum. It is also hypothesized that this event is harmonically related to the tachometer signal. By finding the phase relationship between the event and the tachometer signal, another means of fault classification is provided.

Chapter 3:

Classification of Fault Condition

Following a visual inspection of computer plots, the information extracted in the digital signal processing phase of the research seemed statistically significant. In other words, the variance between fault conditions appeared to be sufficient for classification of gearbox health. To verify this hypothesis, it was necessary to develop a classification algorithm that could take the processed data and return an output corresponding to the condition of the gearbox. In order to do this, separate artificial neural networks were constructed to classify data from each of the eight sensors.

3.1 The Artificial Neural Network

An artificial neural network (ANN) is a machine learning algorithm that can learn a specific task from examples. ANN's are used in pattern and sequence recognition problems where a relationship between problem and solution is known, but not enough is known explicitly to write a program that can relate the two. Essentially, a neural network is a computer model of the human brain. Like a neuron (Fig. 3.1), the processing elements (PE's) have many input paths (dendrites) and a single output path (axon) which is related to a sum of the inputs. These processing elements are interconnected through what is called a "hidden layer," in which various

Figure 3.1: Human Neuron [8]
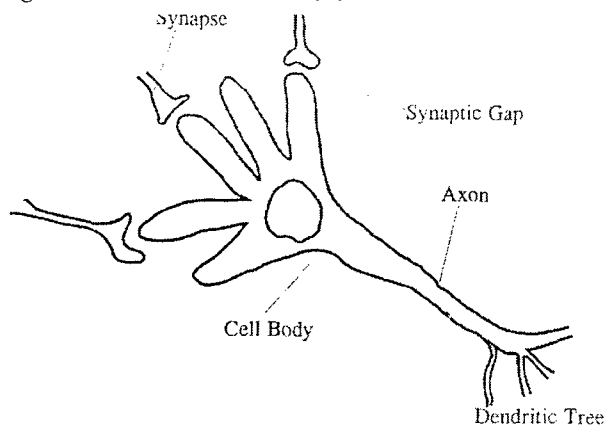
weights are distributed between the connections of the separate processing elements. After all of the information has been re-weighted within the hidden layer, it is directed to the output buffer [7]. One of the most popular network architectures is called the Multi-Layer Perceptron (MLP), Fig 3.3. Figure 3.2 illustrates a typical network structure [3].

Figure 3.2: Structure typical of a neural network [3]
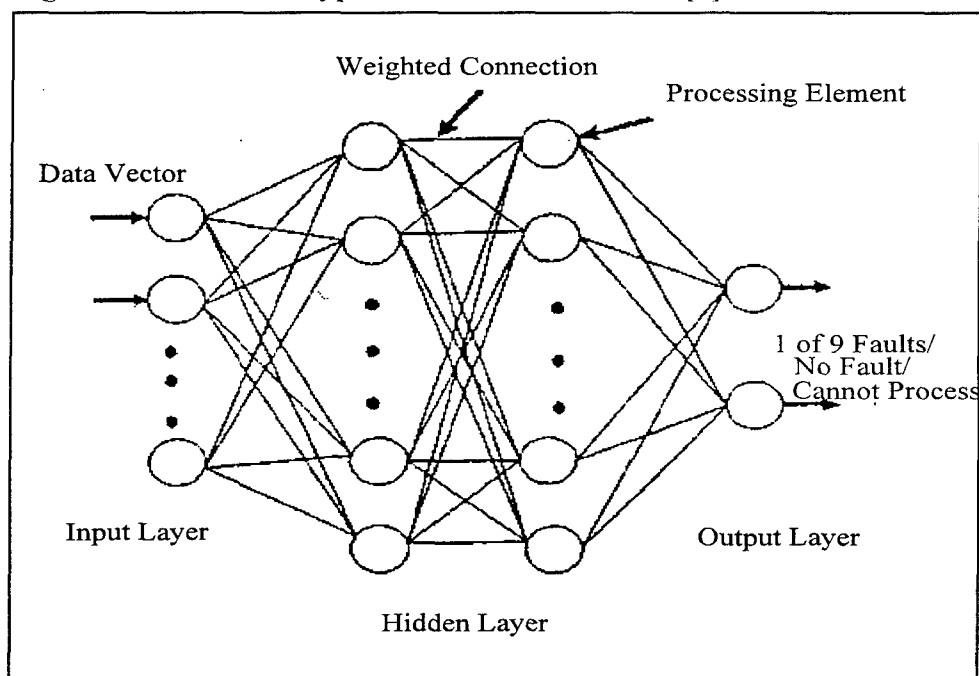
There are two modes of operation for an artificial neural network: learning and recall. In the learning phase, the network is given a training set for which the input and output are known. The neural net then adapts and modifies its connection weights until the output corresponds with the given input. In the recall phase of operation, the network is fed with information not included in the training set. The output is then matched to the most similar training set vector [7].
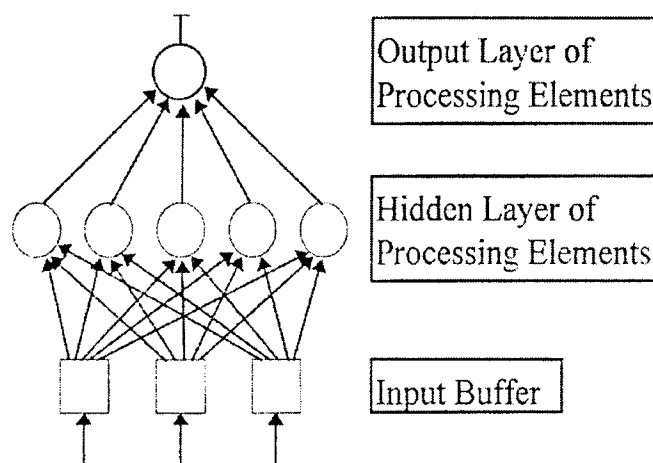
Figure 3.3: Multi-Layer Perceptron [8]

3.2 Construction of the Neural Classifier

The program 'Predict' by NeuralWare Inc. was used to build the fault classification networks. The training method used was based on a technique called gradient back-propagation. Back-propagation involves assigning responsibility for mismatches in classification to each of the processing elements in a network. This re-weighting of connections among the hidden layer is accomplished by propagating the

gradient of the objective function back through the network [8]. The weight update is accomplished via the gradient-descent method as used for simple perceptrons with differentiable units [8].

$\mathbf{x}^{(k)}$=input   (pattern vector)

$\mathbf{d}^{(k)}$=ouput (fault condition represented numerically)

Back-propagation involves two phases of data flow for a given input-output pair ($\mathbf{x}^{(k)}$, $\mathbf{d}^{(k)}$). First, the input pattern is propagated from the input to the output layer in order to form an output $\mathbf{y}^{(k)}$. The difference between $\mathbf{d}^{(k)}$ and $\mathbf{y}^{(k)}$ results in an error signal which is then back-propagated through the previous layers in order to update their connection weights. In order to demonstrate this learning rule, consider a three layer network that consists of m PE's in the input layer, l PE's in the hidden layer, and n PE's in the output layer [6]. A PE q in the hidden layer receives an input of

$$net_q = \sum_{j=1}^{m} v_{qj} * x_j \tag{3.1}$$

and results in an output of

$$z_q = a(net_q) = a(\sum_{j=1}^{m} v_{qj}x_j). \tag{3.2}$$

Therefore, the output for PE i in the output layer is

$$net_i = \sum_{q=1}^{l} w_{iq}z_q = \sum_{q=1}^{l} w_{iq}a(\sum_{j=1}^{m} v_{qj}x_j), \tag{3.3}$$

and it produces a final output of $a(net_i)$.

Next, we must consider the output signals and their back propagation.

$$Error(w) = .5 \sum_{i=1}^{n} [d_i - a(\sum_{q=1}^{l} w_{iq}z_q)]^2 \qquad (3.4)$$

According to the gradient descent method, the connection weights between the hidden

and output layer are then adjusted according to

$$\Delta w_{iq} = -\frac{\eta \partial E}{\partial w_{iq}} \qquad (3.5)$$

Substitution from equations (3.1)-(3.4) and application of the chain rule results in the

equality

$$\Delta w_{iq} = -\eta[\frac{\partial E}{\partial y_i}][\frac{\partial y_i}{\partial net_i}][\frac{\partial net_i}{\partial w_{iq}}] = \eta[d_i - y_i][a'(net_i)][z_q] = \eta \delta_{oi}z_q \quad (3.6)$$

where δ_{oi} is the error signal and its double subscript indicates the ith node in the output

layer. For the weight update between input to hidden connections, the chain rule

coupled with the gradient-descent method is again employed in order to find Δv_{ql} and

δ_{hq}.[6]

$$\delta_{hq} = -\frac{\partial E}{\partial net_q} = -[\frac{\partial E}{\partial z_q}][\frac{\partial z_q}{\partial net_q}] = a'(net_q) \sum_{i=1}^{n} \delta_{oi}w_{iq} \qquad (3.7)$$

The learning rule employed in this project was based on back propagation, and

is called adaptive gradient learning.

3.3 Formation of the Pattern Vector

While attempting to optimize the networks, the most important consideration

involved the construction of statistically significant pattern vectors. Inputs to the

neural network classifier consisted of sampled amplitudes of frequency from the filtered magnitude, envelope, PM, and FM spectra, continua, as well as the RMS values corresponding to the associated signals. Since the record length for each FFT was 32,768 points, it was not feasible to enter all of this information into the network. The sheer amount of data would overwhelm the network and result in a memory error. In order to select only rugged signal features (those not due to noise), the moving average filter technique was used to calculate a signal-to-noise estimate for each point in the magnitude, envelope, PM, and FM spectrum.

Since the data were collected at multiple torque conditions, within-class (same sensor) variance needs to be reduced in order to form a reliable pattern vector. In order to accomplish this, rugged signal features associated with each fault at full torque were extracted using the moving average filter technique. Only the points with a value above a user defined threshold were retained for formation of the pattern vector. By extracting the rugged points corresponding to each fault condition, a 'template' was formed that was guaranteed to contain features common to every fault condition. Since the vectors created for classification were sensor dependent, eight separate classification networks were created. Table 3.1 shows the number of points from the magnitude, envelope, FM, and PM spectra used in the formation of the pattern vectors for each of the eight networks. In addition to the spectral information, the root mean square value of the corresponding signals and the frequency continuum were used as inputs.

Table 3.1: Number of rugged spectral points used as inputs to each network

Spectrum	Net 1	Net 2	Net 3	Net 4	Net 5	Net 6	Net 7	Net 8
Magnitude	159	133	208	178	49	44	131	49
Envelope	82	52	102	73	11	1	75	12
PM	46	13	33	31	1	2	59	10
FM	44	11	31	29	7	3	57	7

3.4 Network Architecture

Following the formation of statistically significant pattern vectors, robust

networks could be created to classify data from each of the eight individual sensors.

The data were split into training and test sets by a ratio of 70/30. The Predict

software itself was responsible for determining the number of input, hidden, and

output processing elements in each neural network.

While many training schemes involve a fixed architecture for the network to

be trained, the software used in this research employed a dynamic method, called

"cascade learning," to determine a suitable number of hidden nodes. This

constructive method was developed by Scott Fahlman of Carnigie Mellon University,

and is characterized by the following properties [8]:

- Hidden PE's are added to the network one at a time during training
- New hidden PE's are connected to both the input buffer and the previously established hidden nodes
- Network construction is stopped when performance shows no further improvement

The software accomplished this by finding the best correlation score during testing and training by modifying connection weights among processing elements. Network architecture for each of the eight networks is shown in Table 3.2. Since the networks were responsible for classifying nine different fault conditions or no-fault, each network consisted of 10 outputs.

Table 3.2 Network Architecture

PE's	Net 1	Net 2	Net 3	Net 4	Net 5	Net 6	Net 7	Net 8
Input	12	8	8	7	8	9	8	9
Hidden	1	2	0	1	2	1	0	0
Output	10	10	10	10	10	10	10	10

3.5 Testing the Classification Networks

Upon construction of several networks and experimenting with the learning parameters governing the training process, eight robust networks were developed (one for each of the eight sensors). The classification results shown in Table 3.3 were obtained.

Table 3.3 Classification Results

Fault	Net 1	Net 2	Net 3	Net 4	Net 5	Net 6	Net 7	Net 8
CGC	90%	90%	90%	90%	90%	80%	100%	90%
HIGC	100%	90%	100%	90%	100%	100%	90%	100%

HIPC2	100%	100%	90%	100%	100%	100%	100%	100%
IPBC1	90%	90%	90%	80%	90%	100%	90%	80%
IPBC2	90%	80%	90%	100%	100%	90%	100%	100%
PBC2	100%	100%	100%	100%	100%	100%	100%	100%
QSC	90%	90%	80%	90%	90%	80%	90%	90%
SBIPS1	100%	90%	100%	100%	90%	100%	100%	100%
SBIPS2	90%	100%	90%	90%	100%	100%	90%	90%
No Fault	100%	90%	100%	100%	100%	100%	90%	100%

These results were collected using test data that was drawn from a separate bank that had not been introduced to the network during the training process. The data bank consisted of ten records for each fault taken at random torque levels. As evident from Table 3.3, the networks performed accurate fault classification with an average accuracy of 94.5% per sensor . By combining the outputs of these eight networks and taking a majority rule. the chance of inaccurate detection of a specific fault is on the order of 10^{-5}. It is also important to note that the fault detector was very good at general fault detection, and was not prone to false warning. In other words, its ability to distinguish a faulted gearbox from a non-faulted gearbox (without specifically identifying the fault) approached 98% on average per sensor. The chances of a misclassification involving gearbox health on a majority of the sensors is on the order of 10^{-7}.

It was initially hypothesized that certain sensors would outperform others based on sensor location. For example, sensor four was located close to the starboard

quill shaft, therefore it was originally believed that sensor four would outperform the other sensors in detecting a cracked quill shaft. This was not found to be the case, because all sensors performed well at detecting any of the faults. Only a minor variance in classification rates among the different sensors was realized. Therefore, the use of all sensors in classification will improve accurate detection probabilities. For simplification in constructing a fault detector and cost savings reasons, a very accurate detector can be implemented using only sensors one, five, and seven. Only when the three sensors agree on a specific fault condition will the detector send a warning. This limits the chances of a false warning to nearly zero, and maintains a detection accuracy of over 90 percent, based on the test data.

Chapter 4:

Overview

The development of a non-invasive fault detector of this type vastly improves the Navy's ability to perform condition based maintenance (CBM) on fleet assets, such as rotary-winged aircraft. Most non-invasive techniques currently in use have trouble identifying a healthy gearbox from a faulted one, much less have the ability to distinguish between specific fault areas. For example, chip detection and oil analysis programs cannot identify gear faults due to root bending fatigue or crack propagation through the gear web, vice through the gear tooth (see Fig. 4.1) [4]. These two detection methods perform well only when a fault results in foreign material being scattered inside the gearbox.

Figure 4.2: Crack Propagation Through Gear Web (NAWC) [4]

The classifier developed in this project not only has the ability to pinpoint faults, but identifies faults due to root bending fatigue and crack propagation. It also has the ability to distinguish fault severity, for example input pinion bearing corrosion (first and second defect levels), and spiral bevel input pinion spalling (first and second defect levels). Another attractive feature of this classifier is its ability to be implemented on existing aircraft using commercial-off-the-shelf

components. With current computers capable of calculating large FFT's in microseconds, the digital signal processing algorithms implemented in this research can be done almost instantaneously, allowing the detector to run in real-time during aircraft operation. By coupling this detection scheme with other procedures such as oil analysis, chip detection, and temperature analysis, a very accurate and reliable fault detector can be implemented at low cost.

In conclusion, pattern recognition through the use of artificial neural networks is a very reliable method for implementing condition based maintenance, and it is a viable and safe alternative to current procedures. Preventative maintenance is costly, and it decreases mission readiness by temporarily grounding usable helicopters. Non-invasive detection of fault conditions will save time and prove cost-effective in both manpower and materials.

Future Work

There are several areas where the classification scheme presented in this work could be improved. Primarily, a relationship between the continuum in the frequency domain and a short duration spike in the time domain is hypothesized to exist. The phase relationship between this event and the tachometer pulse can provide another means to classify fault condition accurately. In other words, the frequency of occurrence of this event as well as its envelope can provide insight into gearbox condition. Due to the complexity involved in the relationship of the tachometer signal to the rotation of specific gearbox parts, time did not permit full investigation of this phenomenon. Although the phase relationship of this event was not accounted for in the pattern vectors used in this project, the continuum associated with this event was used in the primary training sets.

Another extension of this project involves testing the classification networks with data taken from other helicopter gearboxes, such as the SH-60 main gearbox. It was intended for the fault detector developed in this project to be a general detector: in other words, it would be able to classify faults in other rotating machinery sharing similar components (following frequency normalization). This quality could not be tested due to the lack of raw data available.

By combining this method of fault classification with others, such as chip detectors, oil analysis, and component cards, a very reliable system for fault classification can be developed.

Works Cited

[1]   Bengio, Yoshua, Neural Networks for Speech and Sequence Recognition, London: Hodgson Williams Associates, 1995.

[2]   Cameron, B.G., Final Report on CH-46 Aft Transmission Seeded Fault Testing, Westland Helicopters Ltd. Final Technical Report for Office of Naval Research, Contract N00014-92-C-0106, Sept 1993.

[3]   Haas, David J., Lance Flitter, and Joel Milano, Prediction of Helicopter Component Loads Using Neural Networks, AIAA Structures, Structural Dynamics, and Materials Conference 34, 19 April 1993.

[4]   Hardman, Bill, Andrew J. Hess, and Chris Neubert, "SH-60 Helicopter Integrated Diagnostic System (HIDS) Program Experience and Results of Seeded Fault Testing," RTO Meeting Proceedings 7, May 1998, Brussels, Belgium.

[5]   Jenkins, Gwiln and Donald Watt, Spectral Analysis and its Applications, California: Holden-Day, 1968.

[6]   Lee, C.S. George, and Chin-Teng Lin, Neural Fuzzy Systems, Upper Saddle River, NJ: Prentice-Hall, 1996.

[7]   NeuralWare, Inc, Neural Computing, NeuralWare Publishing, 1991.

[8]   NeuralWare, Inc. NeuralWorks Predict, NeuralWare Publishing, 1997.

[9]   Parker, B.E. Jr., H.V. Poor, M.P. Carley, R.J. Ryan, M.J. Szabo, J.W. Staymen, N.A. Nigro, P.D. Mooney, Helicopter Transmission Diagnostics Using Vibration Signature Analysis, Barron Associates, Inc. Final Technical Report for Office of Naval Research, Contract N00014-93-C-0074, July 1995.

[10]   Peebles, Peyton, Probability, Random Variables, and Random Signal Principles, New York: McGraw-Hill, Inc, 1993.

[11]   Sarkady, Dr. Antal A., and Dr. William I. Clement, Preliminary Results of Artificial Neural Network Study of CH-46 Vibration Data, Letter to Lawrence Ash, 25 March, 1995.

[12]   Siegler, J.A., Motor Current Signal Analysis For Diagnosis of Fault Conditions in Shipboard Equipment, U.S.N.A. Trident Scholar Project Report; No 220. 1994.

[13]   Stanley, William D., Digital Signal Processing, Reston, VA: Prentice Hall, 1975.

[14]   Stremler, Ferrel, Introduction to Communication Systems, Massachusetts: Addison-Wesley Publishing Company, 1990.

[15]   United States Navy, OPNAVINST 4790.2G Volume V, Standard Operating Procedures, February, 1998.